Trapping and characterization of the reaction intermediate in cyclodextrin glycosyltransferase by use of activated substrates and a mutant enzyme

Mosi, Renee; He, Shouming; Uitdehaag, Joost; Dijkstra, Bauke W.; Withers, Stephen G.

Published in:
Biochemistry

DOI:
10.1021/bi970618u

Publication date:
1997

Trapping and Characterization of the Reaction Intermediate in Cyclodextrin Glycosyltransferase by Use of Activated Substrates and a Mutant Enzyme†

Renee Mosi,‡ Shouming He,‡ Joost Uitdehaag,§ Bauke W. Dijkstra,§ and Stephen G. Withers*,‡

Department of Chemistry, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z1, and BIOSON Research Institute and Laboratory of Biophysical Chemistry, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands

Received March 18, 1997; Revised Manuscript Received June 2, 1997⊗

ABSTRACT: Cyclodextrin glycosyltransferases (CGTases) catalyze the degradation of starch into linear or cyclic oligosaccharides via a glycosyl transfer reaction occurring with retention of anomeric configuration. They are also shown to catalyze the coupling of maltooligosaccharyl fluorides. Reaction is thought to proceed via a double-displacement mechanism involving a covalent glycosyl−enzyme intermediate. This intermediate can be trapped by use of 4-deoxymaltotriosyl α-fluoride (4DG3αF). This substrate contains a good leaving group, fluoride, thus facilitating formation of the intermediate, but cannot undergo the transglycosylation step since the nucleophilic hydroxyl group at the 4-position is missing. When 4DG3αF was reacted with wild-type CGTase (*Bacillus circulans* 251), it was found to be a slow substrate (k_{cat} = 2 s^{-1}) compared with the parent glycosyl fluoride, maltotriosyl α-fluoride (k_{cat} = 275 s^{-1}). Unfortunately, a competing hydrolysis reaction reduces the lifetime of the intermediate precluding its trapping and identification. However, when 4DG3αF was used in the presence of the presumed acid/base catalyst mutant Glu257Gln, the intermediate could be trapped and analyzed because the first step remained fast while the second step was further slowed (k_{cat} = 0.6 s^{-1}). Two glycosylated peptides were identified in a proteolytic digest of the inhibited enzyme by means of neutral loss tandem mass spectrometry. Edman sequencing of these labeled peptides allowed identification of Asp229 as the catalytic nucleophile and provided evidence for a covalent intermediate in CGTase. Asp229 is found to be conserved in all members of the family 13 glycosyl transferases.

Cyclodextrin glycosyltransferases (CGTase)[1] degrade starch into linear or cyclic oligosaccharides by four distinct reactions. The main one is a cyclization reaction which generates a mixture of cyclic products, referred to as α-, *β*-, or *γ*-cyclodextrins (six, seven, or eight glucose residues, respectively). Other reactions catalyzed include disproportionation, linearization, and hydrolysis. Reaction in each case proceeds with retention of anomeric configuration and is believed to involve a two-step double-displacement mechanism (Scheme 1) (*1*). The first step involves the attack of the catalytic nucleophile at the anomeric center of the sugar with general acid assistance to aid in the leaving group departure. This generates a glycosyl−enzyme intermediate which can then undergo transglycosylation or hydrolysis in a second step with general base assistance to facilitate attack by the incoming group. Both steps proceed via transition states with substantial oxocarbenium ion character. In addition to the two active site carboxylates functioning as the nucleophile and general acid/base catalyst, a third conserved carboxylate is believed to function by modulating the ionization state of the other catalytic residues and producing a more favorable electronic environment for stabilization of the positively charged transition state. The *β*-CGTase from *Bacillus circulans* 251 is currently used in the large scale production of *β*-cyclodextrin, a product which has numerous industrial applications (*2*). Consequently, a detailed investigation of the three-dimensional structure of the enzyme and the mechanism of its transglycosylation reaction has been undertaken in an effort to improve yields and specificity.

Extensive progress has been made in recent years in elucidating the identity and function of the catalytic carboxylates in many α-glucosidases by means of kinetic analysis, site-directed mutagenesis, and X-ray crystallography (see refs 3 and 4 for useful reviews). The nucleotide sequence and 2.0 Å crystal structure of *B. circulans* CGTase have been reported (*5*). The 75 kDa enzyme contains 686 amino acids folded into five domains, labeled A−E. Domains A−C are, together, structurally similar to the α-amylases. The active site of CGTase is located at the N-terminal side of the $(\beta/\alpha)_8$-barrel of domain A. The E-domain has been implicated in starch binding (*6*) and is held in position with respect to domains A−C through an interconnecting region of *β*-sheet (domain D). On the basis of a structure of the enzyme complexed with acarbose (*7*), a potent

† This work was supported by funds from the Natural Science and Engineering Council of Canada.

* Author to whom all correspondence should be addressed (fax, 604-822-2847; e-mail, withers@chem.ubc.ca).

‡ University of British Columbia.

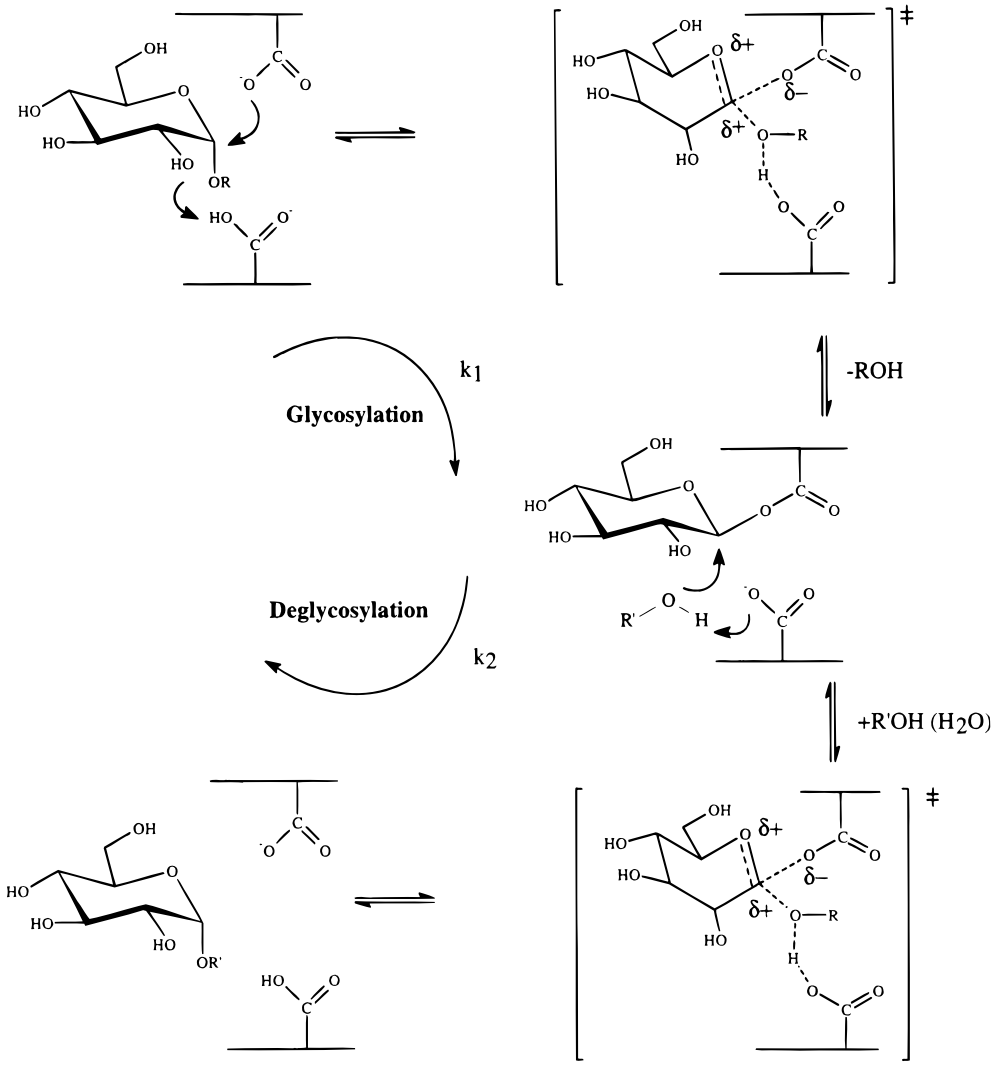
§ University of Groningen.

⊗ Abstract published in *Advance ACS Abstracts,* July 15, 1997.

[1] Abbreviations: CGTase, cyclodextrin glycosyltransferase; α-, *β*-, and *γ*-CD, α-, *β*-, and *γ*-cyclodextrin; 4DG3αF, 4-deoxymaltotriosyl α-fluoride; 4DG2αF, 4-deoxymaltosyl α-fluoride; 4DG3, 4-deoxymaltotriose; 4DG2, 4-deoxymaltose; 4DG4, 4-deoxymaltotetraose; G3αF, maltotriosyl α-fluoride; G2αF, maltosyl α-fluoride; GαF, glucosyl-α-fluoride; G1, glucose; G2, maltose; G3, maltotriose; G4, maltotetraose; G6, maltohexaose; HPLC, high-performance liquid chromatography; LC/MS, liquid chromatography/mass spectrometry; MS/MS, tandem electrospray mass spectrometry; TIC, total ion chromatogram; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis; wt, wild-type.

S0006-2960(97)00618-1 CCC: $14.00

Scheme 1: Proposed Mechanism for a Retaining α-Glycosyltransferase

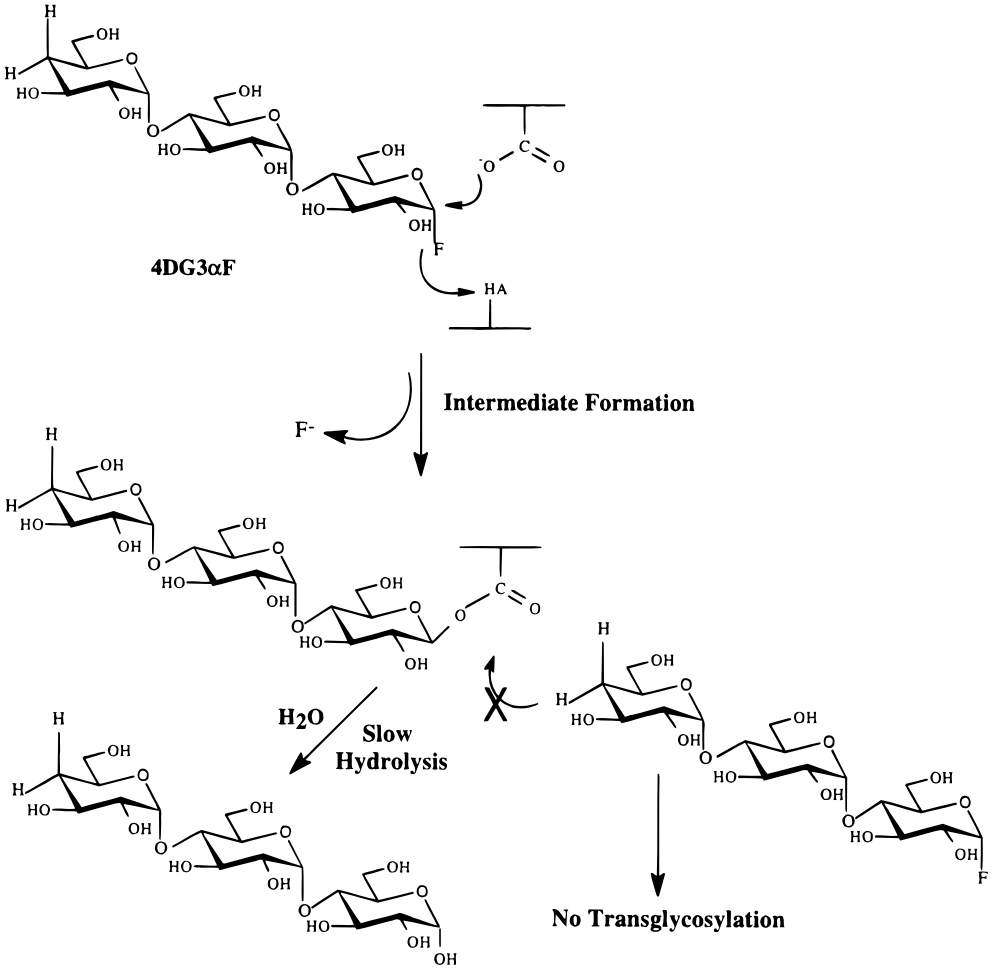


inhibitor of α-amylases [K_i = 10 μM for pig pancreatic α-amylase (*8*)] and CGTases [K_i = 100 nM (*9*)], and of complexes with natural substrates such as maltotetraose and α-cyclodextrin (*8*, *9*) as well as a complex with a maltononaose inhibitor (*10*), the catalytic residues have been identified as Glu257, Asp229, and Asp328 (*B. circulans* 251 numbering). In addition, sequence alignments with other members of the α-amylase family (family 13) have shown that these three carboxylates are completely conserved within the consensus regions and align with other carboxylates which have already been identified as the catalytic residues (*3*). From these results, Glu257 was proposed to be the acid/base catalyst, Asp229 the catalytic nucleophile, and Asp328 the third stabilizing residue as noted earlier. By means of site-directed mutagenesis, three mutants, Asp229Asn, Glu257Gln, and Asp328Asn, have been constructed and studied (*9*). All showed drastically reduced activities using natural substrates. In order to gain insight into the mode of substrate binding a double mutant, Asp229Asn/Glu257Gln, was constructed and crystallized in the presence of α-cyclodextrin and its three-dimensional structure solved (*9*). This structure showed Asp229 to be at a distance of 2.8 Å from the C-1 atom, thus close enough to act as the nucleophile, while Glu257 and Asp328 are somewhat farther away. A similar picture of substrate binding has been shown from structural studies on α-amylases (*8*, *11*−*13*). Definitive identification and confirmation of the covalent nature of the intermediate however require that a reaction intermediate be trapped and characterized.

2-Deoxy-2-fluoroglycosides have proved to be excellent mechanism-based inactivators for retaining β-glycosidases, functioning through the formation of a stabilized glycosyl−enzyme intermediate (*14*−*16*). Unfortunately, this class of compounds has not proven to be useful for studying the mechanisms of retaining α-glucosidases and glycosyltransferases since they do not form a stable intermediate. Indeed, 2-deoxy-2-fluoromaltosyl α-fluoride was found to act as a slow substrate of α-amylase and glycogen-debranching enzyme rather than as an inhibitor (*17*). Apparently the deglycosylation step for α-glycosyltransferases is not slowed sufficiently by substitution of fluorine at C-2 to allow accumulation of an intermediate.

An alternative strategy for trapping the intermediate was therefore adopted for glycogen-debranching enzyme in which a substrate was synthesized which could undergo the first step, formation of the glycosyl−enzyme, but not the second, transglycosylation to a second molecule of substrate. This was achieved by synthesis of 4-deoxymaltotriosyl α-fluoride (4DG3αF), a substrate which contains a good leaving group, fluoride, thereby facilitating formation of an intermediate. However, the absence of a suitably disposed nucleophile to attack that intermediate on the second substrate molecule,

Scheme 2: Mechanism for the Accumulation of an Intermediate on CGTase in the Presence of 4DG3αF and Subsequent Turnover via Hydrolysis



as a result of deoxygenation at C-4, resulted in trapping of the intermediate thereby allowing subsequent identification of the attachment site (*18*).

The aim of this study was to apply a similar approach to the labeling and identification of the catalytic nucleophile of CGTase. It was therefore necessary first to determine whether maltotriosyl α-fluoride could act as a satisfactory substrate for *B. circulans* CGTase. This seemed likely since other glycosyl fluorides had been shown to act as good donors for CGTases in enzymatic synthetic reactions (*19*). If this were the case then a glycosyl fluoride substrate which was deoxygenated at the 4-position of the nonreducing end (such as 4DG3αF) should still undergo the first step (glycosylation) but not the second step (deglycosylation) resulting in accumulation of an intermediate as shown in Scheme 2. Such studies are described in this paper.

MATERIALS AND METHODS

General. Wild-type, Glu257Gln, and Asp229Asn CGTases were purified according to the previously published method of Knegtel et al. (*9*). G3αF, G2αF, and GαF were synthesized according to the method of Junnemann et al. (*20*). 4DG3αF and 4DG2αF were synthesized as described previously by Lindhorst et al. (*21*). Pepsin was obtained from Boehringer Mannheim. All other chemicals were obtained from Sigma Chemical Co. and were used without further purification.

Kinetic Analysis. The reactions of CGTase with the glycosyl fluoride analogues were monitored by following the release of fluoride ion using an Orion ion selective fluoride electrode. In a typical experiment, an appropriate concentration of glycosyl fluoride in 50 mM citrate buffer (pH 6.0) was warmed to 30 °C for 10 min and the reaction initiated by the addition of enzyme [final concentrations: wild-type (0.25 μg/mL), Glu257Gln (24 μg/mL), Asp229Asn (50 μg/mL)]. The reaction was typically followed for approximately 10 min or until 10% of the starting material was consumed. Rates were determined at 6−8 substrate concentrations ranging from $^1/_5$ to 5 times the estimated K_m value. From the experimental rate versus substrate concentration data, the values of K_m and k_{cat} were calculated using the program GraFit (*22*).

HPLC Analysis of Reaction Products. To analyze the reaction products of wild-type and mutant CGTases, each enzyme (at identical concentrations to those used in the kinetic analysis) was incubated at room temperature with the substrate of interest (at concentrations of 5 × K_m) in 50 mM citrate buffer, pH 6.0. At specified times, 20 μL aliquots were removed and injected directly onto the HPLC. HPLC analysis was performed using a Dynamax 60 Å NH_2 (4.6 × 250 mm, mean particle size 5 μm) column from Rainin Instrument Co. (Woburn, Ma) linked to a refractive index detector. A mobile phase of 60/40 CH_3CN/H_2O at a flow rate of 0.8 mL/min was used. Whenever possible, retention

times were matched to authentic standards which were run under identical conditions.

MS Analysis of Reaction Products. Mass spectra were recorded on a Sciex API 300 mass spectrometer (Sciex, Thornhill, Ontario). In each case, the reaction mixture was injected onto a Dynamax NH_2 (60 Å) HPLC column as specified above, running with a mobile phase of 65/35 CH_3CN/H_2O (to ensure maximum separation of peaks) at 0.8 mL/min. A postcolumn splitter allowed the introduction of 15% of the reaction mixture into the mass spectrometer at 50 μL/min and permitted the addition of 50/50 $MeOH/H_2O$ (containing 5 mM ammonium acetate) at 5 μL/min as an ionizer. Whenever possible, the masses of authentic standards were obtained through direct flow injection into the mass spectrometer.

Labeling and Proteolysis of CGTase. Labeling of CGTase was performed by incubating Glu257Gln (10 μL, 6 mg/mL) with an excess of 4DG3αF (5 μL, 125 mM) for 10 min in 50 mM citrate buffer, pH 6.0. Proteolysis was achieved by first lowering the pH through the addition of 15 μL of 50 mM sodium phosphate buffer (pH 2) followed by immediate addition of 10 μL of pepsin (0.4 mg/mL in pH 2 buffer), and the mixture was incubated at room temperature for 1 h. ES-MS analysis and SDS-PAGE of the proteolytic digests confirmed that CGTase was completely digested under these conditions.

Electrospray Mass Spectrometry. The analyses of the protein and peptide samples were carried out using a Sciex API-300 mass spectrometer interfaced with a Michrom UMA HPLC system (Michrom Bioresources, Inc., Auburn, CA). Intact CGTase [10−20 μg, wild-type, Glu257Gln (unlabeled or labeled)] was introduced into the mass spectrometer through a microbore PLRP column (1 × 50 mm) and eluted with a gradient of 20−100% solvent B at a flow rate of 50 μL/min over 5 min (solvent A, 0.06% trifluoroacetic acid, 2% acetonitrile in water; solvent B, 0.05% trifluoroacetic acid, 90% acetonitrile in water). The MS was scanned over a range of 400−2000 Da with a step size of 0.5 Da and a dwell time of 1 ms. The peptides were analyzed by loading a 10 μL sample of the pepsin digest (4.5 mg/mL) onto a C18 column (Relasil, 1 × 150 mm) and eluting at a flow rate of 50 μL/min with a gradient of 0−60% B over 60 min. The proteolytic mixture was first examined in LC/MS mode and then in neutral loss mode.

In the single-quadrupole (normal LC/MS) mode, the MS conditions were as follows: the mass analyzer was scanned over the range of 300−2400 Da with a step size of 0.5 Da, a dwell time of 1 ms, an ion source voltage (ISV) of 4.8 kV, and an orifice energy (OR) of 50 V. The neutral loss spectra were obtained in the triple-quadrupole mode searching for the loss of *m/z* 235.5 (4DG3αF) or 154.5 (4DG2αF) which corresponds to the loss of the inhibitor label from a peptide which is doubly charged. A scan range of 400−2200 amu was used with a step size of 0.5 amu and a dwell time of 1 ms. Other parameters are as follows: ISV = 5 kV, OR = 45 V.

Chemical Sequencing. The peptic digest (120 μL of 1.5 mg/mL) of the 4DG3-labeled CGTase was separated by HPLC using a Deltapak 300 Å C-18 column (3.9 × 150 mm, mean particle size 15 μm) from Waters eluted with a gradient of 0−60% B over 60 min at 700 μL/min. A 13:1 postcolumn splitter allowed the introduction into the mass spectrometer of a portion of the peptides to be checked for purity, and the fractions containing the 4DG3-labeled peptide were collected for sequencing. Prior to solid phase sequence analysis, peptides (2−5 μg in 40% CH_3CN/0.5% TFA) were coupled to arylamine-functionalized poly(vinylidene fluoride) membranes (Sequelon AA, Milligen/Bioresearch) using *N*-ethyl-*N*'-[3-(dimethylamino)propyl] carbodiimide. Peptide sequences were determined both by solid phase Edman degradation on a Milligen/Bioresearch model 6600 protein sequenator using standard protocols and by on-line HPLC analysis of the resulting phenylthiohydantoins.

Table 1: Kinetic Constants Determined for the Reaction of *B. circulans* CGTase (Wild-Type and Mutants) with Glycosyl Fluorides

substrate	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m (mM^{-1} s^{-1})
		Wild-Type	
G3αF	2.5	275	111
G2αF	10.9	314	28.8
GαF	32.5	333	10.2
4DG3αF	0.073	2	27.5
		Glu257Gln	
G3αF	0.37	1.0	2.70
4DG3αF	0.027	0.6	22.2
		Asp229Asn	
G3αF	0.10	0.14	1.4

RESULTS AND DISCUSSION

α-*Glycosyl Fluoride Substrates with wt CGTase*. Maltotriosyl α-fluoride proved to be a good substrate for CGTase as shown in Table 1 (k_{cat} = 275 s^{-1} and K_m = 2.5 mM). This compares very well to a k_{cat} of 350 s^{-1} for the natural substrate, starch (*9*). Thus, although CGTase cannot normally use a trisaccharide as a substrate, the presence of the fluoride at the anomeric position as a good leaving group is sufficient to replace the transition state stabilization ordinarily provided by interactions between the enzyme and the oligosaccharide leaving group. Thus it serves as an activated substrate and eliminates the need for a longer oligosaccharide.

HPLC analysis of the reaction products revealed that the dominant reaction occurring is transglycosylation to produce longer maltooligosaccharyl fluorides, followed by cyclization. The major products were β-CD and a mixture of primarily longer oligosaccharides arising from a variety of cyclization, coupling, and disproportionation reactions (Figure 1A). Thus, the maltotriosyl α-fluoride acts as a glycosyl donor, forming a maltotriosyl−enzyme intermediate which then reacts with another equivalent of maltotriosyl fluoride to form maltohexaosyl fluoride and so on.

Two other glycosyl fluorides were also investigated as substrates for *B. circulans* CGTase. As seen in Table 1, both glucosyl fluoride (GαF) and maltosyl fluoride (G2αF) served as substrates, but with reduced efficiencies compared to G3αF, as seen in k_{cat}/K_m values. Interestingly, the k_{cat} values for the three substrates were similar indicating that, once bound, there was no particular advantage in a longer oligosaccharide. The progressive decrease in k_{cat}/K_m arose from changes in K_m values. While values most likely reflect an increase in affinity due to the presence of the additional sugar residues, they could also, to some extent, reflect an increase in the rate of the glycosylation step relative to deglycosylation as sugar residues are added, leading to accumulation of the intermediate with the consequence of reduced K_m values. All substrates, after an extended period of time, gave β-CD as the major product as seen in Figure

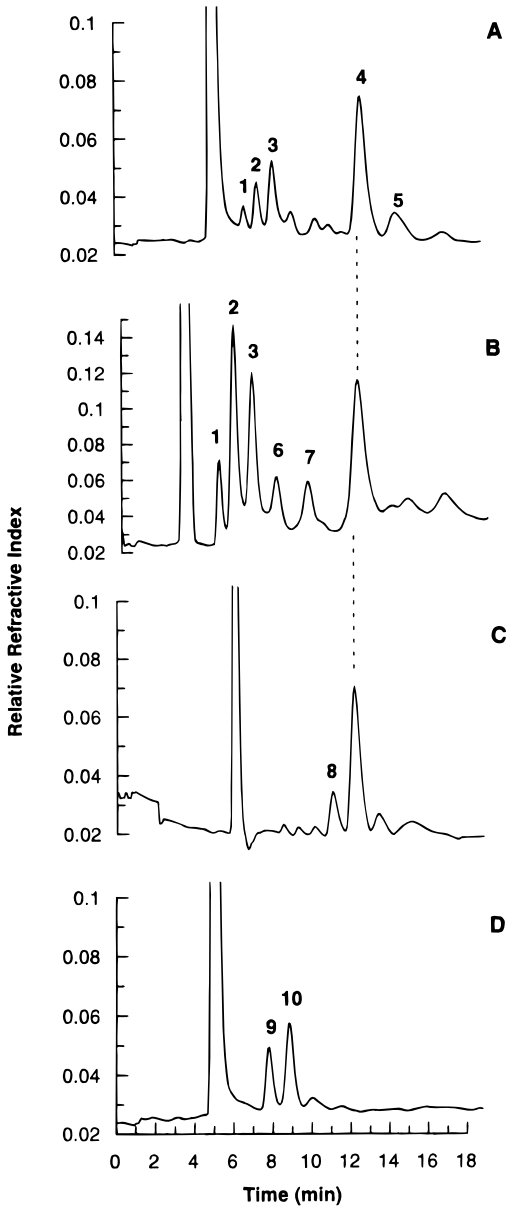


FIGURE 1: HPLC analysis of the products of the reaction of wild-type CGTase with **(**A**)** G3αF after 1 h reaction time, **(**B**)** G2αF after 1 h reaction time, **(**C**)** GαF after 1 h reaction time, and **(**D**)** 4DG3αF after 1 h reaction time (1 = GαF, 2 = G2αF, 3 = G, 4 = β-CD, 5 = γ-CD, 6 = G2, 7 = G3, 8 = α-CD, 9 = G/4DG2, 10 = G2/4DG3).

1B for the reaction of CGTase with G2αF and Figure 1C for the reaction of CGTase with GαF.

Reaction of CGTase with 4-deoxymaltotriosyl α-fluoride (4DG3αF) was investigated, and as seen in Table 1, 4DG3αF was found to act as a substrate, albeit slow, indicating that the enzyme intermediate is formed but turns over at a significant rate. The mechanism of how this turnover occurs was resolved by HPLC analysis of reaction products as seen in Figure 1D. 4-Deoxymaltotriose (4DG3) was the major product, resulting from simple hydrolysis of the maltotriosyl−enzyme intermediate. In addition, however, small amounts of the disproportionation products, G1, G2, and 4DG2, were observed arising from cleavage within the oligosaccharide moiety (Figure 1D). Kinetic analysis gave a K_m value of 0.073 mM, which is approximately 35 times lower than that for G3αF (Table 1). The low K_m value is almost certainly due to the deglycosylation step (k_2) being very much slower than glycosylation (k_1), thus accumulation of an intermediate rather than inherent differences in true dissociation constants. This would arise as a consequence of removal of the glycosyl transfer pathway, as is reflected in the lower k_{cat} value (reduced over 100-fold) for 4DG3αF. Unfortunately, the relatively short half-life ($t_{1/2}$ = 0.5 s) of the intermediate precluded any investigation by mass spectrometry, necessitating some additional measures to increase its lifetime.

Investigation of Glu257Gln CGTase. A solution to the problem of the still relatively fast hydrolysis of the accumulating intermediate lay in the presumed acid/base catalyst mutant Glu257Gln. In previous work with cellulases and β-glucosidase (*23*−*25*), it had been shown that by removing the acid/base catalyst and using an activated substrate such as a glycosyl fluoride with a good leaving group the intermediate could be readily accumulated. Since the leaving group does not need acid catalysis the first step is still fast. However, since that residue also serves as the general base catalyst for the second step, deglycosylation is slowed, allowing an accumulation of the intermediate. Thus, it was believed that the combination of the presumed acid/base catalyst mutant of CGTase, Glu257Gln, with an incompetent substrate such as 4DG3αF should provide a means to accumulate an intermediate of longer lifetime.

G3αF as a Substrate for Glu257Gln. G3αF was found to be a substrate for Glu257Gln CGTase as shown in Table 1, but with significantly reduced k_{cat} and K_m values. The k_{cat} was reduced almost 300-fold, and K_m was reduced 7-fold compared with wild-type; thus k_{cat}/K_m was reduced 40-fold.

HPLC analysis of products showed a slightly different pattern to that seen with the wild-type enzyme. Although β-CD was the main cyclization product obtained, considerably more enzyme-catalyzed hydrolysis occurred, producing maltotriose. In addition, a lower percentage of disproportionation products was obtained (Figure 2A).

4DG3αF as an Incompetent Substrate for Glu257Gln. The K_m for 4DG3αF was 3-fold lower in comparison with wild-type and over 10-fold when compared with the reaction of Glu257Gln CGTase and G3αF, likely indicating a greater accumulation of intermediate. The k_{cat} value was reduced by a factor of 3.3-fold over the reaction with wild-type indicating that the contribution of this group to the general base-catalyzed hydrolysis pathway is of the same order of magnitude as that for the reaction of Glu257Gln with G3αF, consistent with the idea that the reaction is occurring primarily via hydrolysis. This very small rate decrease as a consequence of the removal of the general base catalyst is surprising. We are currently investigating the reasons for this, and the possibility exists of an alternative reaction mechanism under these conditions. The half-life ($t_{1/2}$) of the intermediate was estimated to be approximately 1.1 s from the k_{cat} value. Interestingly, HPLC analysis of the products of the reaction of Glu257Gln CGTase with 4DG3αF showed only 4DG3 as a major product (Figure 2B), and essentially no disproportionation products were seen as with wild-type CGTase and 4DG3αF (Figure 1D) even after overnight incubation.

Mass Spectrometric Evidence for a Covalent Enzyme−*Substrate Intermediate.* Upon incubation of Glu257Gln CGTase with an excess of 4DG3αF, an increase in molecular weight equivalent to a 4DG3 moiety on the intact protein [(74 513 + 471) ± 3 Da] was observed by ES-MS. The

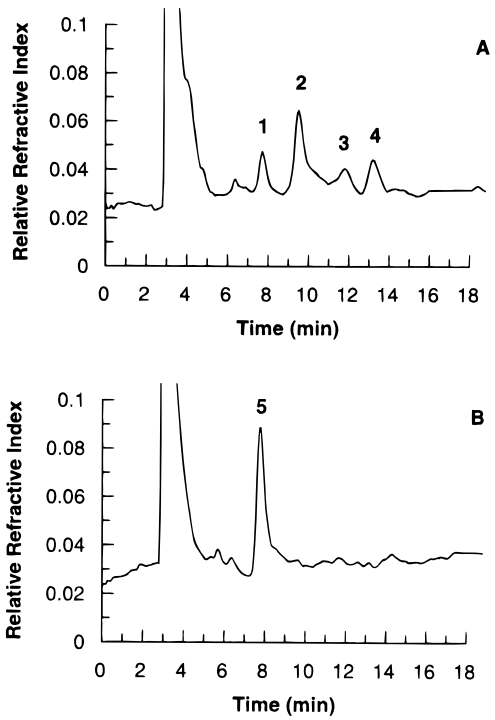


FIGURE 2: HPLC analysis of the products of the reaction of Glu257Gln CGTase with **(**A**)** G3αF after 1 h reaction time and **(**B**)** 4DG3αF after 1 h reaction time (1 = G2, 2 = G3, 3 = G4, 4 = β-CD, 5 = 4DG3).

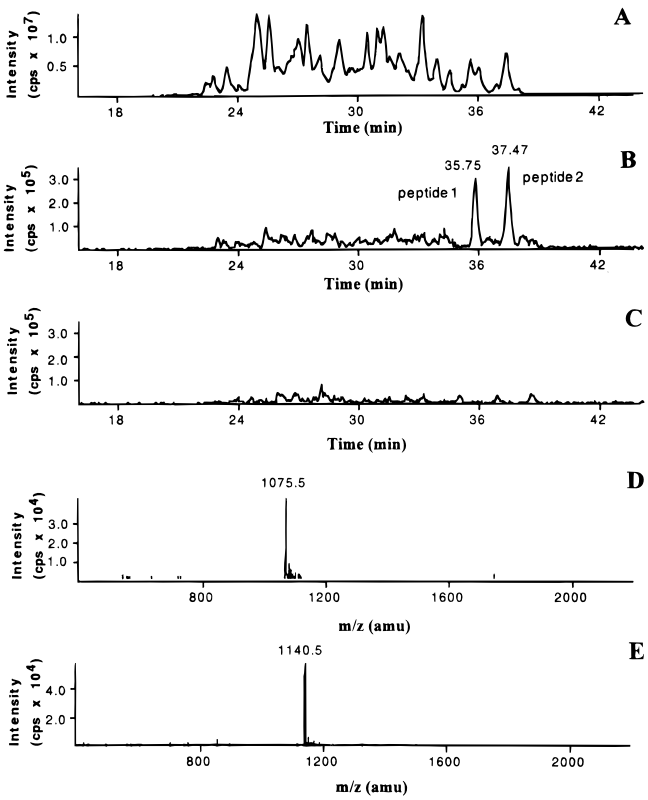


FIGURE 3: Electrospray mass spectrometry experiments on Glu257Gln CGTase proteolytic digests: **(**A**)** labeled with 4DG3αF, total ion chromatogram in normal MS mode; **(**B**)** labeled with 4DG3αF, TIC in neutral loss mode; **(**C**)** unlabeled, in neutral loss mode; **(**D**)** mass spectrum of peptide 1 in panel B; **(**E**)** mass spectrum of peptide 2 in panel B.

equivalent disaccharide, 4DG2αF also acted as an incompetent substrate, and similar ES-MS analysis of Glu257Gln CGTase incubated with 4DG2αF revealed a mass increase of 309 ± 3 Da consistent with the formation of a 4DG2− enzyme complex.

Identification of the Peptide Containing the Nucleophile. Identification of the attachment site of the 4-deoxymaltotriosyl label was achieved by using the neutral loss tandem MS method (*26*). In this technique, ions are subjected to limited fragmentation by collisions with an inert gas (nitrogen) in a collision cell located between the two mass analyzers Q1 and Q3. Because the ester linkage between the inhibitor and the peptide is one of the more labile linkages present, facile homolytic cleavage of this bond occurs resulting in the loss of the sugar as a neutral species, with the peptide retaining its original charge. The mass spectrometer is then set so that Q1 and Q3 are scanned in a linked manner such that only those ions that lose the mass of the lost sugar moiety can pass through both quadrupoles and be detected.

The 4DG3-labeled Glu257Gln mutant and a control (not exposed to 4DG3αF) were digested using pepsin. The resulting peptide mixtures were separated by reverse phase HPLC using the ES-MS as a detector, revealing a large number of peptides (Figure 3A). The peptide with the 4DG3 label was located by MS/MS analysis using the neutral loss experiment. The spectrometer was first scanned in the neutral loss mode searching for a peptide which loses a neutral species of 471 Da, which corresponds to the mass of the 4-deoxymaltotriosyl label. However, no significant peaks were detected. When the spectrometer was scanned for the mass loss of *m/z* 235.5, corresponding to the loss of a sugar from a doubly charged peptide, two strong peaks were observed, labeled peptide 1 and 2 (Figure 3B). These two peaks, having retention times of 35.75 and 37.5 min and *m/z* values of 1075.5 and 1140.5 Da, respectively, were not seen in the control digest (Figure 3C) and presumably represent two different 4-deoxymaltotriosyl-labeled peptides. Because the doubly charged labeled peptides have *m/z* values of 1075.5 Da (Figure 3D) and 1140.5 Da (Figure 3E), the singly charged unlabeled peptides must have masses of 1679 [(2 × 1075.5) − 2 − 471 + 1 H] and 1809 [(2 × 1140.5) − 2 − 471 + 1 H]. A computer-generated search of all possible peptides in the Glu257Gln CGTase mutant yielded 18 possible peptides with molecular weight 1679 ± 2 Da and 34 possible peptides with molecular weight of 1809 ± 2 Da. Of these, seven sets of peptides contained overlapping amino acid sequences. It is assumed that the two peptides arise from different modes of proteolysis of the same labeled protein since only a single oligosaccharide was initially attached.

When CGTase was incubated with 4DG2αF, digested, and analyzed by neutral loss, the same two peptides were found to be labeled (data not shown). Because of the large size of the labeled peptides, their definitive identification by MS/MS was precluded since good fragmentation was not achieved. However, as can be seen in Figure 3A, the peptides were readily separable by HPLC and after purification, were subjected to Edman sequencing. As shown in Table 2, sequences of ^{229}DAVKHMPFGWQKSF^{242} for peptide 1 and ^{229}DAVKHMPFGWQKSFM^{243} for peptide 2 were obtained, clearly indicating, as expected, that they are overlapping peptides. The catalytic nucleophile must be an amino acid which is common to both peptides and by analogy with other glycosidases and glycosyltransferases is either a glutamate or aspartate (*3, 16, 27*). The only carboxylic

Table 2: Edman Degradation of Glu257Gln CGTase Active Site Peptides 1 and 2[a]

peptide 1				peptide 2			
cycle	PTH derivative	yield (pmol)	cDNA	cycle	PTH derivative	yield (pmol)	cDNA
1	Asp	42	Asp	1	Asp	43	Asp
2	Ala	60	Ala	2	Ala	32	Ala
3	Val	91	Val	3	Val	44	Val
4	Lys	20	Lys	4	Lys	10	Lys
5	His	3	His	5	His	2	His
6	Met	24	Met	6	Met	7	Met
7	Pro	9	Pro	7	Pro	2	Pro
8	Phe	19	Phe	8	Phe	4	Phe
9	Gly	20	Gly	9	Gly	2	Gly
10	Trp	19	Trp	10	Trp	3	Trp
11	Gln	6	Gln	11	Gln	5	Gln
12	Lys	2	Lys	12	Lys	2	Lys
13	Ser	2	Ser	13	Ser	Q	Ser
14	Phe	Q	Phe	14	Phe	Q	Phe
				15	Met	Q	Met

[a] Q = qualitative assignment.

amino acid contained within these sequences is Asp229 suggesting that this is indeed the catalytic nucleophile.

Analysis of Nucleophile Mutant. In order to further probe the role of the catalytic nucleophile, a kinetic analysis of the Asp229Asn CGTase mutant from *B. circulans* was performed. Values of k_{cat} and K_m for G3αF were found to be reduced 2000- and 30-fold by comparison with wild-type CGTase. This compares with a reduction of 300- and 7-fold, respectively for the Glu257Gln mutant when assayed with the same substrate. When Asp229Asn was assayed for β-cyclodextrin-forming activity with starch as a substrate, a decrease of the same order of magnitude [(2.3×10^4)-fold] was observed compared with a decrease of 4×10^3 times for Glu257Gln (*9*). A similar decrease in activity (10^3 drop from wild-type) was observed for an Asp229Ala mutant from *B. circulans* CGTase strain 8 (*28*). In addition, similar reductions in activity have been observed with the analogous nucleophile mutants in other family 13 glycosylhydrolases. For example, when assayed for β-cyclodextrin-forming activity, the mutant Asp229Asn from an alkalophilic *Bacillus* showed at least a 10^3-fold drop in activity compared with wild-type (*29*). In addition, when the equivalent residue to Asp229 (Asp176) in *Bacillus subtilis* α-amylase was mutated to Asn, a 10^5-fold decrease in activity against starch was measured (*30*). A similar dramatic loss of activity was seen for the equivalent residue in neopullulanase (*31*) and Taka α-amylase (*32*), although these two proteins were not purified before measurement. In contrast, β-retaining glycosidases generally exhibit a much larger decrease in activity when the nucleophile is mutated. For example, mutation of the nucleophile Glu358 to alanine in *Agrobacterium faecalis* β-glucosidase resulted in a 10^6-fold reduction in activity (*33*).

It is unlikely that the residual activity exhibited by the Asp229Asn mutant is due to a wild-type contaminant because its measured K_m is so low (0.1 mM) compared with that of wild-type (2.5 mM). Instead, it has been speculated from crystallographic studies that the residual activity exhibited by the nucleophile mutant may arise from the third active site carboxylate (Asp328) being in a position to partially take over for the nucleophile (*9*). Indeed, the low K_m value might well indicate that an intermediate was accumulating. However, mass spectral analysis of Asp229Asn CGTase incubated with either G3αF or 4DG3αF did not provide any indication of an intermediate accumulating on the enzyme. In addition, it is unlikely that the low level of activity exhibited by Asp229Asn is due to the presence of a contaminating glycosyltransferase or glycosidase as HPLC analysis showed that α- and β- cyclodextrins were among the reaction products, thus suggesting a CGTase activity (data not shown).

In addition to defining the role of Asp229 as the catalytic nucleophile, these results also help to clarify the available X-ray structural data on the functions of the three conserved carboxylates in CGTase. In native CGTase, a strong hydrogen bond (2.5 Å) exists between Glu257 and Asp328 (*5*, *28*), indicating that one of the carboxylates is most likely protonated, most probably Glu257. In the structure with acarbose bound (*7*), the hydrogen bond between Asp328 and Glu257 is broken, and the Glu257 side chain moves to within hydrogen-bonding distance of the glycosidic oxygen of the scissile bond. Thus, it was suspected that Glu257 is the proton donor that initiates the reaction. Further, from the same structure of CGTase complexed with acarbose, Asp229 is 2.8 Å from the C-1 atom and is thus close enough to act as either a nucleophile or an acid/base catalyst. A similar scenario is observed for the X-ray structure of porcine α-amylase complexed with acarbose (*8*). Here, Asp197 (the equivalent residue to Asp229) lies on the opposite side of the inhibitor-binding cleft to the other two key carboxylates, Glu233 and Asp300, at a distance of 3.3 Å from the anomeric carbon C-1 of the inhibitor center, compared to a distance of 3.5 Å for Glu233. Hence, the positions of Asp197 and Glu233 allow either of them to be the nucleophilic partner of the reaction as they are both properly oriented to produce a β-linked glycosyl−enzyme intermediate. However, the presence of a hydrogen bond between Glu233 and the glycosidic nitrogen linking the A and B residues of acarbose suggests that it is more likely that Glu233 acts as the acid/base catalyst in the reaction. Movement of the Glu257 side chain was also observed when a crystal of Asp229Ala from *B. circulans* strain 8 was briefly exposed to β-cyclodextrin (*28*).

CONCLUSIONS

Confirmation of the identity of the nucleophile of *B. circulans* CGTase 251 as Asp229 was achieved using a unique combination of site-directed mutagenesis, an incompetent substrate, and mass spectrometry. This work has also provided the first concrete evidence for a covalent enzyme−substrate intermediate for *B. circulans* 251 CGTase. In two other cases evidence of a covalent glycosyl−enzyme intermediate in family 13 glycosidases/glycosyltransferases has been obtained from rapid trapping studies with natural substrates. Low-temperature ^{13}C NMR experiments have given evidence for the formation of a β-carboxylacetal ester covalent adduct between maltotetraose and porcine pancreas α-amylase (*34*). Other evidence for a stabilized intermediate comes from experiments with *Streptococcus sobrinus* α-glycosyltransferase where an aspartic acid was identified as the catalytic nucleophile by rapidly denaturing a reaction mixture of enzyme and radiolabeled sucrose and isolating the labeled peptide (*27, 35*). In one other family 13 enzyme, yeast α-glucosidase, two mechanism-based glycosidase inactivators, 5-fluoro-α-D-glucosyl fluoride and 5-fluoro-β-L-idosyl fluoride, were used in conjunction with mass spectrometry to identify the catalytic nucleophile as Asp214 (the homolo-

gous residue to Asp229), thereby also confirming the existence of a covalent intermediate (*36*).

Future studies include the possibility of using 4DG3αF in crystallographic studies of the covalent intermediate bound to CGTase. This should help to provide valuable information toward understanding the mechanism of reaction of CGTase at the catalytic site. In turn, an improved understanding of the processes occurring at the reaction site will help to clarify the contributions of other residues to substrate binding and product specificity. Such information could be useful in creating enzymes which produce specific types of cyclodextrin products.

ACKNOWLEDGMENT

We thank Dr. Curtis Braun and Dr. Thisbe Lindhorst for helpful discussions regarding the synthesis of G3αF and 4DG3αF. Thanks to Dr. Dirk Penninga and Prof. Lubbert Dijkhuizen for generous gifts of preparations of mutant CGTases.

REFERENCES

1. Koshland, D. E. (1953) *Biol. Rev. 28*, 416−436.
2. Schmid, G. (1989) *Trends Biotechnol. 7*, 224−248.
3. Svensson, B. (1994) *Plant Mol. Biol. 25*, 141−157.
4. Davies, G., & Henrissat, B. (1995) *Structure 3*, 853−859.
5. Lawson, C. L., van Montfort, R., Strokopytov, B., Rozeboom, H. J., Kalk, K. H., de Vries, G. E., Penninga, D., Dijkhuizen, L., & Dijkstra, B. W. (1994) *J. Mol. Biol. 236*, 590−600.
6. Penninga, D., van der Veen, B. A., Knegtel, R. M., van Hijum, S. A., Rozeboom, H. J., Kalk, K., Dijkstra, B., & Dijkhuizen, L. (1996) *J. Biol. Chem. 271* (51), 32777−32784.
7. Strokopytov, B., Penninga, D., Rozeboom, H. J., Kalk, K. H., Dijkhuizen, L., & Dijkstra, B. W. (1995) *Biochemistry 34*, 2234−2240.
8. Qian, M., Haser, R., Buisson, G., Duee, E., & Payan, F. (1994) *Biochemistry 33*, 6284−6294.
9. Knegtel, R. M., Strokopytov, B., Penninga, D., Faber, O. G., Rozeboom, H. J., Kalk, K. H., Dijkhuizen, L., & Dijkstra, B. W. (1995) *J. Biol. Chem. 270*, 29256−29264.
10. Strokopytov, B., Knegtel, R. M. A., Penninga, D., Rozeboom, H. J., Kalk, K. H., Dijkhuizen, L., & Dijkstra, B. W. (1996) *Biochemistry 35*, 4241−4249.
11. Buisson, G., Duee, E., Haser, R., & Payan, F. (1987) *Embo J. 6*, 3909−3916.
12. Larson, S. B., Greenwood, A., Cascio, D., Day, J., & McPherson, A. (1994) *J. Mol. Biol. 235*, 1560−1584.
13. Qian, M., Haser, R., & Payan, F. (1993) *J. Mol. Biol. 231*, 785−799.
14. Withers, S. G., Rupitz, K., & Street, I. P. (1988) *J. Biol. Chem. 263*, 7929−7932.
15. Street, I. P., Kempton, J. B., & Withers, S. G. (1992) *Biochemistry 31*, 9970−9978.
16. Withers, S. G., & Aebersold, R. (1995) *Protein Sci. 4*, 361−372.
17. Braun, C. (1995) Ph.D. Thesis, University of British Columbia, Vancouver, BC, Canada.
18. Braun, C., Lindhorst, T., Madsen, N. B., & Withers, S. G. (1996) *Biochemistry 35*, 5458−5463.
19. Cottaz, S., Apparu, C., & Driguez, H. (1991) *J. Chem. Soc., Perkin. Trans. I* 2235−2241.
20. Junnemann, J., Thiem, J., & Pedersen, C. (1993) *Carbohydr. Res. 249*, 91−94.
21. Lindhorst, T. K., Braun, C., & Withers, S. G. (1995) *Carbohydr. Res. 268*, 93−106.
22. Leatherbarrow, R. J. (1990) *Gra-Fit, Version 2.0;* Erithacus Software Ltd., Staines, U.K.
23. MacLeod, A., Tull, D., Rupitz, K., Warren, R. A. J., & Withers, S. G. (1996) *Biochemistry 35*, 13165−13172.
24. Wang, Q., Trimbur, D., Graham, R., Warren, R. A., & Withers, S. G. (1995) *Biochemistry 34*, 14554−14562.
25. Gebler, J. C., Trimbur, D. E., Warren, A. J., Aebersold, R., Namchuk, M., & Withers, S. G. (1995) *Biochemistry 34*, 14547−14553.
26. Miao, S., Ziser, L., Aebersold, R., & Withers, S. G. (1994) *Biochemistry 33*, 7027−7032.
27. Mooser, G., Hefta, S. A., Paxton, R. J., Shively, J. E., & Lee, T. D. (1991) *J. Biol. Chem. 266*, 8916−8922.
28. Klein, C., Hollender, J., Bender, H., & Schulz, G. E. (1992) *Biochemistry 31*, 8740−8746.
29. Nakamura, A., Haga, K., Ogawa, S., Kuwano, K., Kimura, K., & Yamane, K. (1992) *FEBS 296*, 37−40.
30. Takase, K., Matsumoto, T., Mizuno, H., & Yamane, K. (1992) *Biochim. Biophys. Acta 1120*, 281−288.
31. Kuriki, T., Takata, H., Okada, S., & Imanaka, T. (1994) *J. Bacteriol. 173*, 6147−6152.
32. Nagashima, T., Toda, S., Kitamoto, K., Gomi, K., Kumagai, C., & Yamane, K. (1992) *Biosci. Biotech. Biochem. 56*, 207−210.
33. Wang, Q., Graham, R. W., Trimbur, D., Warren, R. A. J., & Withers, S. G. (1994) *J. Am. Chem. Soc. 116*, 11594−11595.
34. Tao, B. Y., Reilly, P. J., & Robyt, J. F. (1989) *Biochim. Biophys. Acta 995*, 214−220.
35. Mooser, G. (1992) In *The Enzymes* (Sigman, D. S., Ed.) pp 187−233, Academic Press Inc., New York.
36. McCarter, J. D., & Withers, S. G. (1996) *J. Am. Chem. Soc. 118*, 241.

BI970618U